BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 2: e25525
doi: 10.3897/biss.2.25525

OPEN ACCESS

Conference Abstract

Establishment of the ARCTOS-GGBN Data Pipeline

Teresa Jegelewicz Mayfield‡, Mariel Campbell‡, Kyndall Hildebrandt§, Carla Cicero|, Dusty McDonald¶, Joseph Cook‡, John Demboski#

‡ Museum of Southwestern Biology, Albuquerque, United States of America
§ University of Alaska Museum, Fairbanks, United States of America
| University of California, Berkeley, Berkeley, United States of America
¶ Arctos, Sacramento, United States of America
# Denver Museum of Nature & Science, Denver, United States of America

Corresponding author: Mariel Campbell (campmlc@unm.edu)

Received: 04 Apr 2018 | Published: 17 May 2018

Citation: Mayfield T, Campbell M, Hildebrandt K, Cicero C, McDonald D, Cook J, Demboski J (2018) Establishment of the ARCTOS-GGBN Data Pipeline. Biodiversity Information Science and Standards 2: e25525. https://doi.org/10.3897/biss.2.25525

Abstract

Arctos (https://arctosdb.org), an online collection management information system, was developed in 1999 to manage museum specimen data and to make those data publicly available. The portal (arctos.database.museum) now serves data on over 3.5 million cataloged specimens from more than 130 collections throughout North America in an instance at the Texas Advanced Computing Center. Arctos also is a community of museum professionals that collaborates on museum best practices and works together to improve Arctos data richness and functionality for on-line museum data streaming. In 2017, three large Arctos genomics collections at the Museum of Southwestern Biology (MSB), Museum of Vertebrate Zoology, Berkeley (MVZ), and University of Alaska Museum of the North (UAM), received support from GGBN to create a pipeline for publishing data from Arctos to the GGBN portal. Modifications to Arctos included standardization of controlled vocabulary for tissues; changes to the data structure and code tables with regard to permit information, container history, part attributes, and sample quality; implementation of interfaces and protocols for parent-child relationships between tissues, tissue subsamples, and DNA extracts; and coordination with the DWC community to ensure that all GGBN data standards and formatting are included in the standard DWC export in order to finalize the pipeline to GGBN. The addition of these three primary Arctos biorepositories to the GGBN network will add over 750,000 tissue and DNA records representing over 11,000 species

© Mayfield T et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and 667 families. These voucher-based archives represent primarily vertebrate taxa, with growing collections of arthropods, endoparasites, and incipient collections of microbiome and environmental samples associated with online media and linked to GenBank and other external databases. The high-quality data in Arctos complement and significantly extend existing GGBN holdings, and the establishment of an Arctos-GGBN pipeline also will facilitate future collaboration between more Arctos collections and GGBN.

Presenting author

Teresa Jegelewicz Mayfield